

- N. B. (1) Question no. 1 compulsory.
(2) Answer any four out of the remaining questions.

Q.1 Attempt

20

- a) What are major issues in data mining?
- b) Explain different OLAP operations.
- c) Difference between database and data warehouse.
- d) Write a short note on Linear regression.

- Q.2 a) Explain constraint based and multilevel association rules with an example. 10
b) Explain market basket analysis and uses of it. 10

- Q.3 a) Explain BIRCH method of clustering with an example. 10
b) Explain Regression. Write short note on Non-linear regression. 10

- Q.4 a) Explain data cleaning, data transformation and Integration with an example. 10
b) Apply Bayesian classification to predict class of new tuple (Nicol, Female, 1.67m), Use the following data. 10

Person ID	Name	Gender	Height	Class
1	Kristina	Female	1.6 m	Short
2	Jim	Male	2 m	Tall
3	Maggie	Female	1.9 m	Medium
4	Martha	Female	1.85 m	Medium
5	John	Male	2.8 m	Tall
6	Bob	Male	1.7 m	Short
7	Clinton	Male	1.8 m	Medium
8	Nyssa	Female	1.6 m	Short
9	Kathy	Female	1.65 m	Short

- Q.5 a) What are outlier. Explain outlier analysis. 10
b) Explain K-means clustering and solve the following with k=3 10
{2,3,6,8,9,12,15,18,22}

- Q.6 a) Explain Business Intelligence issues. 10
b) Describe the steps involved in data mining when viewed as a process of Knowledge discovery. 10

- Q.7 Short note on any **Three** 20
a) Application of Web Mining
b) Market segmentation
c) Sequence Mining in transaction
d) Agglomerative clustering.